

REPORT DOCUMENTATION PAGE			Form Approved OMB NO. 0704-0188		
<p>The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA, 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.</p> <p>PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.</p>					
1. REPORT DATE (DD-MM-YYYY) 20-08-2012		2. REPORT TYPE Conference Proceeding		3. DATES COVERED (From - To) -	
4. TITLE AND SUBTITLE A Bayesian analysis of scale-invariant processes			5a. CONTRACT NUMBER W911NF-12-1-0095		
			5b. GRANT NUMBER		
			5c. PROGRAM ELEMENT NUMBER 611102		
6. AUTHORS Jingfeng Wang, Rafael L. Bras, Veronica Nieves			5d. PROJECT NUMBER		
			5e. TASK NUMBER		
			5f. WORK UNIT NUMBER		
7. PERFORMING ORGANIZATION NAMES AND ADDRESSES Georgia Tech Research Corporation Office of Sponsored Programs Georgia Tech Research Corporation Atlanta, GA 30332 -0420			8. PERFORMING ORGANIZATION REPORT NUMBER		
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Research Office P.O. Box 12211 Research Triangle Park, NC 27709-2211			10. SPONSOR/MONITOR'S ACRONYM(S) ARO		
			11. SPONSOR/MONITOR'S REPORT NUMBER(S) 61346-EV.1		
12. DISTRIBUTION AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.					
13. SUPPLEMENTARY NOTES The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.					
14. ABSTRACT We have demonstrated that the Maximum Entropy (ME) principle in the context of Bayesian probability theory can be used to derive the probability distributions of those processes characterized by its scaling properties including multiscaling moments and geometric mean. We started from a proof-of-concept case of a power-law probability distribution, followed by the general case of multifractality aided by the wavelet representation of the cascade model. The					
15. SUBJECT TERMS A Bayesian Analysis of Scale-Invariant Processes					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT UU	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON Rafael Bras
a. REPORT UU	b. ABSTRACT UU	c. THIS PAGE UU			19b. TELEPHONE NUMBER 404-385-2700

## Report Title

A Bayesian analysis of scale-invariant processes

### ABSTRACT

We have demonstrated that the Maximum Entropy (ME) principle in the context of Bayesian probability theory can be used to derive the probability distributions of those processes characterized by its scaling properties including multiscaling moments and geometric mean. We started from a proof-of-concept case of a power-law probability distribution, followed by the general case of multifractality aided by the wavelet representation of the cascade model. The ME formalism leads to the probability distribution of the multiscaling parameter and those of incremental multifractal processes at different scales. Compared to other methods, the ME method significantly reduces computational cost by leaving out unimportant details. The ME distributions have been evaluated against the empirical histograms derived from the drainage area of river network, soil moisture and topography. This analysis supports the assertion that the ME principle is a universal and unified framework for modeling processes governed by scale-invariant laws. The ME theory opens new possibilities of extracting information of multifractal processes beyond the scales of observation.

**Conference Name:** BAYESIAN INFERENCE AND MAXIMUM ENTROPY METHODS IN SCIENCE AND ENGINEERING: 31

**Conference Date:** July 09, 2011

## A Bayesian analysis of scale-invariant processes

Veronica Nieves, Jingfeng Wang, and Rafael L. Bras

Citation: *AIP Conf. Proc.* **1443**, 56 (2012); doi: 10.1063/1.3703620

View online: <http://dx.doi.org/10.1063/1.3703620>

View Table of Contents: <http://proceedings.aip.org/dbt/dbt.jsp?KEY=APCPCS&Volume=1443&Issue=1>

Published by the American Institute of Physics.

---

### Additional information on AIP Conf. Proc.

Journal Homepage: <http://proceedings.aip.org/>

Journal Information: [http://proceedings.aip.org/about/about\\_the\\_proceedings](http://proceedings.aip.org/about/about_the_proceedings)

Top downloads: [http://proceedings.aip.org/dbt/most\\_downloaded.jsp?KEY=APCPCS](http://proceedings.aip.org/dbt/most_downloaded.jsp?KEY=APCPCS)

Information for Authors: [http://proceedings.aip.org/authors/information\\_for\\_authors](http://proceedings.aip.org/authors/information_for_authors)

### ADVERTISEMENT



*Submit Now*

### Explore AIP's new open-access journal

- Article-level metrics  
now available
- Join the conversation!  
Rate & comment on articles

# A Bayesian Analysis of Scale-Invariant Processes

Veronica Nieves\*, Jingfeng Wang<sup>†</sup> and Rafael L. Bras<sup>†</sup>

*\*Jet Propulsion Laboratory, Pasadena, California, USA.*

*<sup>†</sup>Georgia Institute of Technology, Atlanta, Georgia, USA.*

**Abstract.** We have demonstrated that the Maximum Entropy (ME) principle in the context of Bayesian probability theory can be used to derive the probability distributions of those processes characterized by its scaling properties including multiscaling moments and geometric mean. We started from a proof-of-concept case of a power-law probability distribution, followed by the general case of multifractality aided by the wavelet representation of the cascade model. The ME formalism leads to the probability distribution of the multiscaling parameter and those of incremental multifractal processes at different scales. Compared to other methods, the ME method significantly reduces computational cost by leaving out unimportant details. The ME distributions have been evaluated against the empirical histograms derived from the drainage area of river network, soil moisture and topography. This analysis supports the assertion that the ME principle is a universal and unified framework for modeling processes governed by scale-invariant laws. The ME theory opens new possibilities of extracting information of multifractal processes beyond the scales of observation.

**Keywords:** Bayesian Statistics, Maximum Entropy, Scale-Invariant Laws, Multifractality, Environmental Sciences.

**PACS:** 02.50.-r, 05.45.Df

## INTRODUCTION

This paper reports the investigation of two recent studies [1, 2], motivated by the need to show the practical application of the Maximum Entropy (ME) principle [3, 4, 5] in the description of statistics of multiscaling processes. These processes are considered: 1. Drainage Area of River Network (DARN), 2. Incremental surface soil moisture and topography fields viewed as a Self-similar Processes (SP), 3. Incremental topography as a Multifractal Process (MP). Our goal is to establish the equivalence of both the empirical and ME distributions, yet ensuring computational efficiency. Instead of computing empirical histograms from large amount of data, only some details are relevant to capture the macroscopic behaviour by the Bayesian analysis. Compared to other multifractal models, the ME predicts the probability distribution of multiscaling processes at different scales using fewer parameters. The key to make successful predictions lies in the correct identification of essential physical constraints. Based on previous studies [6, 7, 8, 9, 10], we assume that the geometric structures were an attractive candidate in determining statistics. It is well known that variables such as drainage area, discharge and energy release, soil moisture, increments of topography, etc., are self-similar characterized by the multiscaling moments or power-law distributions. For the case of increments of topography, the multifractal condition is also an important scaling property described by means of the wavelet representation of the cascade model [11, 12, 13, 14, 15, 16, 9, 17, 18, 19]. It should be noted that, even though the physi-

cal principles underlying various multiscaling processes might be different, the statistics are found to be dominated by some geometric properties. The similar patterns that these processes exhibit at different resolutions lead the “most probable” configuration. Our findings also shed light on the information beyond the limited observational scales.

## RESULTS AND DISCUSSION

Here, we first summarize the previous results of the ME distributions of: 1. DARN, 2. SP, and 3. MP. The procedure involved is explained in *Materials and methods*, where the following constraints have been used: i) geometric means, ii) multiscaling moments, and iii) multifractal condition. Further details can be found in [1, 2].

1. The ME distribution for a given DARN,  $a$ , with  $A_1 < a < A_2$ ,  $A_1$  and  $A_2$  as the upper and lower limit of  $a$ , respectively, subject to constraint i),  $A_g$ , is the power-law distribution:

$$p(a) = \left\{ \frac{1 - \mu}{A_2^{1-\mu} - A_1^{1-\mu}} \right\} a^{-\mu}, \quad (1)$$

where for the case of  $A_2 \gg A_1$ , the *Lagrangian multiplier*  $\mu$  is a function of  $A_g$ ,

$$\mu = 1 + \left( \log \frac{A_g}{A_1} \right)^{-1} \quad \text{or} \quad \mu = 1 - \left( \log \frac{A_2}{A_g} \right)^{-1}, \quad (2)$$

depending on whether  $\mu > 1$  or  $\mu < 1$ .

2. Given the incremental process  $|z_1 - z_2|$  of a SP field  $z(\vec{x})$ , where  $\vec{x}$  is a two-dimensional location vector, and  $z_1 \equiv z(\vec{x}_1)$  and  $z_2 \equiv z(\vec{x}_2)$ , subject to i) and ii), the ME joint distribution can be expressed as

$$p(z_1, z_2) = \frac{1}{Z} |z_1 - z_2|^{-\mu_0} \exp \left\{ - \sum_{q=1}^M \mu_q |z_1 - z_2|^q \right\}, \quad (3)$$

where  $\mu_q$  is the *Lagrangian multiplier* associated with the  $q^{\text{th}}$  moment constraint.

3. Denote the incremental process of  $z$  at a location  $\vec{x}$  over a two-dimensional domain by  $\Delta_r z \equiv |z_1^r - z_2^r|$  where  $z_1^r \equiv z(r\vec{x}_1)$  and  $z_2^r \equiv z(r\vec{x}_2)$  are a re-scaled  $z$  by scale factor  $r > 0$ . The ME joint distribution of a MP,  $\Delta_r z$ ,  $p_z(z_1^r, z_2^r)$ , satisfying the three scaling properties (i), ii) and iii)), is given by

$$p_z(z_1^r, z_2^r) = \frac{1}{Z'} (\Delta_{r'} z)^{-\mu_0} \exp \left\{ - \sum_{n=1}^N \lambda_n A_{r/r'}^n - \sum_{m=1}^M \mu_m (\Delta_{r'} z)^m \right\} \quad (4)$$

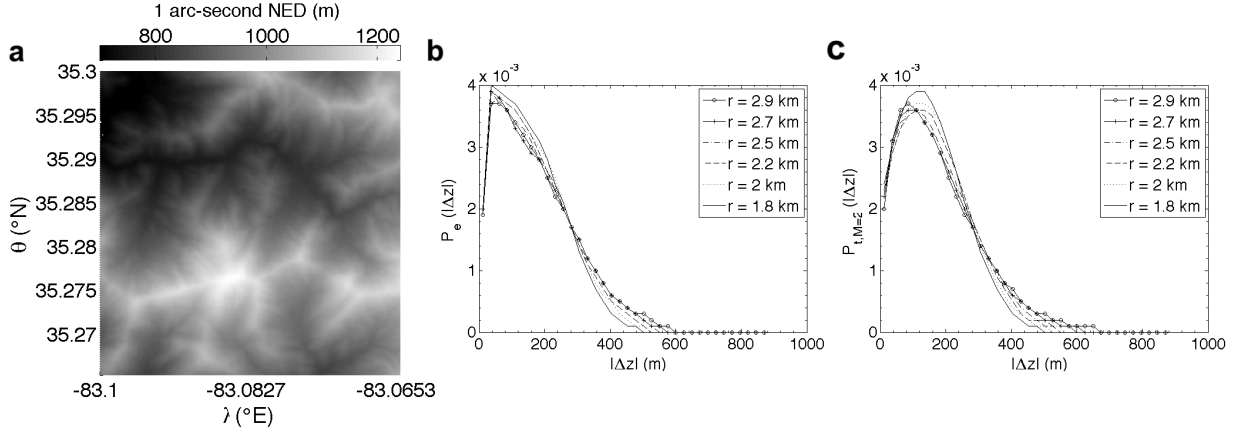
where the Lagrangian multipliers  $\lambda_n$  and  $\mu_m$  are again related to the given constraints, and  $A_{r/r'}$  is a multiscaling parameter (the cascade variable [20]) independent of  $\Delta_r z$  and  $\Delta_{r'} z$  for any pair of  $r$  and  $r'$  with ME distribution as

$$p_A(A_{r/r'}) = \frac{1}{Z^n} \exp \left\{ - \sum_{n=1}^N \lambda_n A_{r/r'}^n \right\}. \quad (5)$$

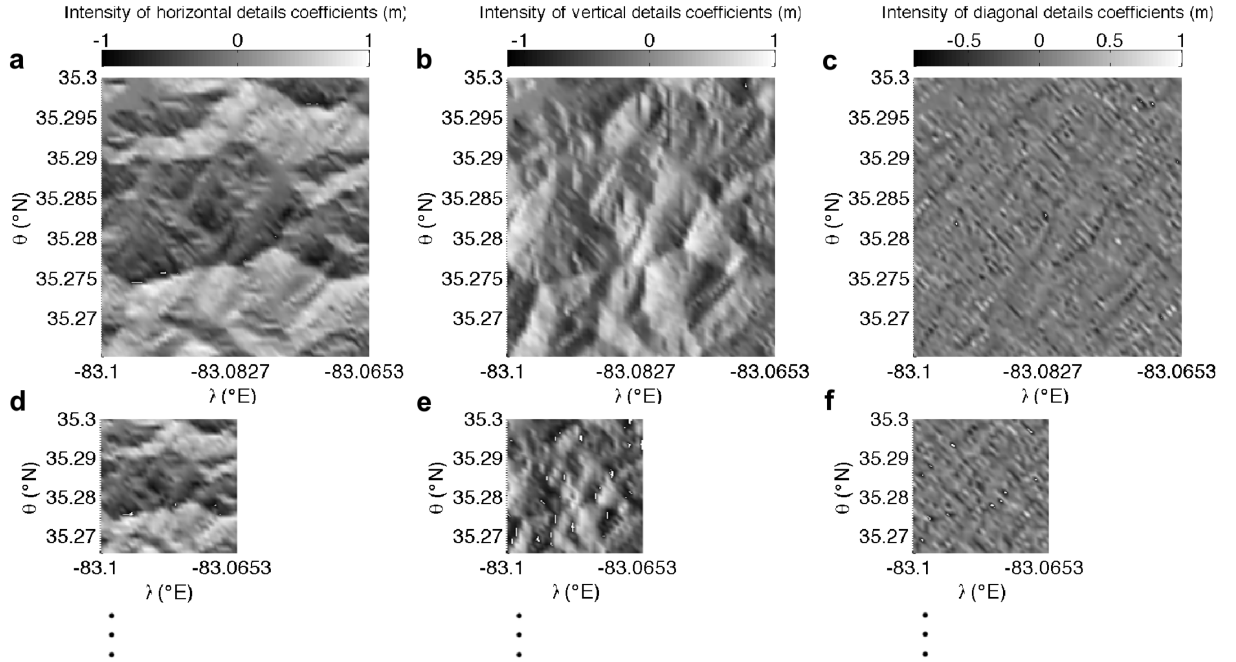
Two data products were used to evaluate the ME distributions of a SP, Eq. (3), against empirical histograms over fifteen regions: the AMSR-E soil moisture L2B product and the USGS National Elevation Dataset (NED). The latter was also used to evaluate the predicted power-law exponent  $\mu$ , Eq. (2), in the Grand Canyon area and some river networks of Puerto Rico, as well as Eqs. (4)- (5) in the characterization of a MP for seven regions. For the case 1, it was found that predicted and empirical results (not shown) were in close agreement, revealing the link between statistics and the macroscopic properties of the system. Similarly, the ME distributions of soil moisture and elevation increments (case 2) were compared with the empirical ones for different separation distances and  $M = 2$ , resulting in an almost identical agreement. Only results for the case of topography are included in this paper (see Fig. (1)). The geographical coordinates of the reported region are [-83.10,-83.04] in east longitude and [35.25,35.3] in north latitude. The same region is shown for the test of case 3 predictions for  $N = 2$  and  $M = 2$ . In this case, the wavelet analysis ([21, 22]) was applied in order to study the distributions at several scales of observation and different directions (see Fig. (2)), where in the wavelet domain:  $z$  becomes  $\alpha_{j,l,\vec{k}}$ ,  $\Delta_r z$  and  $\Delta_{r'} z$  turns into  $\Delta \alpha_{j,l,\vec{k}}$  and  $\Delta \alpha_{j-1,l,\frac{\vec{k}}{2}}$ , respectively, and  $\eta_{j,l,\vec{k}}$  corresponds to  $A_{r/r'}$ . Figures (3)- (4) compare the ME and empirical distributions with log-log regression coefficients of around 0.99 and 0.87, respectively. Higher correlation coefficients are found at any scale for bigger domains. As far as we know, the results are independent of the type of wavelet bases (Haar, Daubechies, Symmlet, Coiflet, Battle-Lemarié) when sufficient wavelet coefficients are included. Only second order moment is considered in this study. A better agreement between empirical and theoretical distributions is expected for higher moments (i.e.,  $M > 2, N > 2$ ).

## CONCLUSIONS

This study re-examines some previously-raised issues concerning the derivation of probability distribution of multiscaling processes using the ME principle in the context of Bayesian probability theory [1, 2]. Typically probability distributions are obtained computing histograms that require a considerable number of data points. The case studies of drainage area of a river network, soil moisture and topography suggest that the ME yields robust macroscopic statistics with tractably reduced set of physical constraints that gives the same result as empirical histograms. Identification of those constraints is often difficult since the behaviour of the complex Earth system is not fully understood. Nonetheless, physical processes following scaling laws manifest some geometric patterns that reveal underlying mechanisms. Common geometric properties associated with scale-invariant processes include: the multifractal condition, the multiscaling moments, and the geometric means. ME translates these physical assumptions into macroscopic predictions. Because the ME discards irrelevant information, it also improves computational efficiency. For a multifractal system such as topography, in addition to the prob-

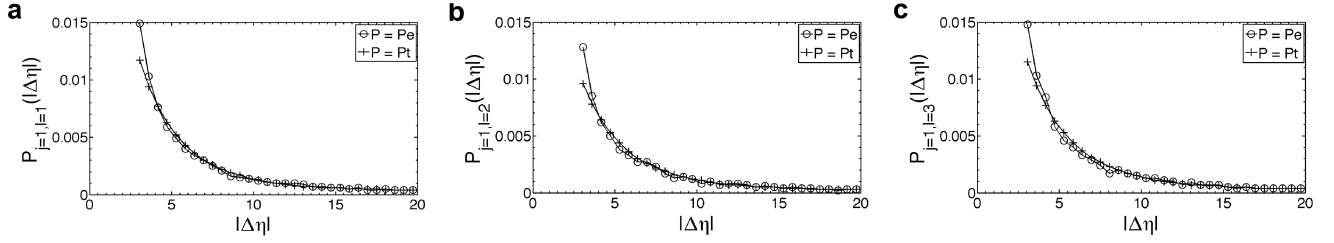


**FIGURE 1.** From left to right: (a) 1 arc second NED map represented in longitude  $\lambda$  and latitude  $\theta$  degrees for a domain of  $128 \times 128$  points. (b) and (c) Empirical ( $P_e$ ) and the MaxEnt distributions for  $M = 2$  ( $P_t$ ) according to Eq. (3). All probabilities are plotted versus the absolute value of the increments  $\Delta z = |z(\vec{x}_1) - z(\vec{x}_2)|$  for different separation distances  $|\vec{x}_1 - \vec{x}_2|$ .

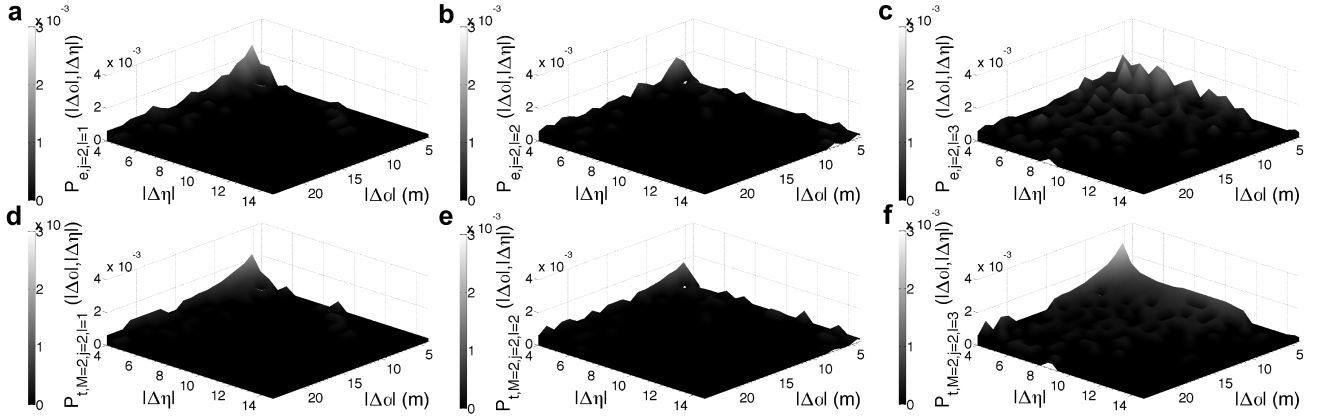


**FIGURE 2.** Wavelet multiresolution representation of the NED map shown in Fig.1 (a) for two levels of decomposition using Haar basis ( $\alpha_{j,l,\vec{k}}$ ). From top to bottom: re-scaled Fig.1 (a) map into  $64 \times 64$  (a, b and c) and  $32 \times 32$  (d, e and f) points domains (i.e.,  $j = 1$  and 2). From left to right: horizontal (a and d), vertical (b and e) and diagonal (c and f) details (i.e.,  $l = 1, 2, 3$ ).

ability distribution of the incremental process, the ME predicts the probability distribution of the multiscaling parameter (the cascade variable) using the wavelet analysis of the cascade model. These distributions are used, for instance, in the study of fluvial



**FIGURE 3.** Empirical ( $P_e$ ) and predicted ME distributions for  $N = 2$  ( $P_t$  or Eq. (5) in the wavelet domain) of the random variable  $\eta_{j,l,\vec{k}}$  (i.e.,  $A_{r,l'}$ ) which characterizes the scaling properties of a multifractal process. According to the wavelet decomposition, **a**, **b** and **c** stand for  $l = 1, 2$  and  $3$ , respectively, from the scale of observation  $j = 1$  to the  $j = 2$  one (downscale cascade). Results correspond to order 3 Battle-Lemarié basis. Correlation coefficients are of around 0.99.



**FIGURE 4.** Probability distributions of increments of topography  $\Delta\alpha_{j-1,l,\frac{\vec{k}}{2}}$  (i.e.,  $\Delta_{r'}z$ ) as a function of  $\eta_{j,l,\vec{k}}$  (i.e.,  $A_{r,l'}$ ) for a given separation distance  $|\vec{x}_1 - \vec{x}_2|$ . From top to bottom: empirical ( $P_e$  for **a**, **b** and **c**) and predicted ME distributions for  $N = 2$  and  $M = 2$  ( $P_t$  or Eq. (4) in the wavelet domain for **d**, **e** and **f**), both at  $j = 2$ . From left to right:  $l = 1, 2$  and  $3$  details, respectively. Results correspond to order 3 Battle-Lemarié basis. Correlation coefficients are of around 0.87.

geomorphology, where the evolution of landscapes clearly plays an important role since it is strongly influenced by the presence of storms, i.e., large floods [13]. The statistics under all possible variations of elevation and changes of resolution can be now described from partial information consistent with the identified constraints. These findings open the possibility of gathering information across the spatial scales outside of those of observations of many research fields. The existence of the cascade model is not confined to this particular context and ME can be applied to other multifractal processes.



## MATERIALS AND METHODS

### Data

The AMSR-E L2B soil moisture product provides estimates of soil moisture of the top 1 cm derived from brightness temperature at 10.7 GHz with resolution of 38 km, re-gridded into a global cylindrical 25 km Equal-Area Scalable Earth Grid (EASE-Grid). The NED raster elevation data of one arc-second resolution (30 m) over the continental US are derived from multiple satellites. NAD83 and NAV88 are consistently used as horizontal and vertical datums, respectively, and all the data are re-gridded using a geographic projection. Soil moisture accuracy (root-mean-square difference) and the vertical accuracy (absolute error) of the NED product is on the order of 7-15 m. More details can be found at: <http://nsidc.org/data/amsre/>, and <http://ned.usgs.gov/>.

### Maximum Entropy Principle

The maximum entropy (ME) principle states that “out of all the possible probability distributions which agree with the given constraint information, select the one that is maximally non-committal with regard to missing information” [5]. The Shannon information entropy  $S_I$  of a probability density function  $p(x)$  of a continuous variable  $x$  is defined as [23, 4, 5],

$$S_I = - \int p(x) \log \left[ \frac{p(x)}{m(x)} \right] dx, \quad (6)$$

where  $m(x)$  is the *ignorance prior* [4].

Suppose that some testable information, (i.e., set of available experimental results or conserved quantities) can be expressed in terms of constraints in the form,

$$\int f_k(x) p(x) dx = F_k, \quad k = 0, \dots, M. \quad (7)$$

where  $f_k$  is an arbitrary function of  $x$ , and  $F_k$  the given parameters.

Maximizing  $S_I$  in Eq. (6) subject to the constraints of Eq. (7) leads to a (normalized) probability density function,

$$p(x) = \frac{m(x)}{Z} \exp \left\{ - \sum_{k=0}^M \mu_k f_k(x) \right\} \quad (8)$$

where  $\mu_k$  is the *Lagrangian multiplier* associated with the constraint  $F_k$  through,

$$F_k = \frac{\partial \log Z}{\partial \mu_k}, \quad (9)$$

and  $Z$  is the well-known *partition function*,

$$Z(\mu_0, \dots, \mu_M) = \int m(x) \exp \left\{ - \sum_{k=0}^M \mu_k f_k(x) \right\} dx. \quad (10)$$

$p(x)$  in Eq. (8) is referred to as the "ME distribution". In this study, we assume  $m(x) = 1$ .

## ACKNOWLEDGMENTS

This work was supported by the U.S. Army RDECOM ARL Army Research Office under Grants No. W911NF-07-1-0126 and No. W911NF-10-1-0236.

## REFERENCES

1. V. Nieves, J. Wang, R. L. Bras, and E. Wood, *Phys Rev Lett* **105**, 118701 (2010).
2. V. Nieves, J. Wang, and R. L. Bras, *Geophys. Res. Lett.* **38**, L17405 (2011).
3. E. T. Jaynes, *Phys. Rev.* **108**, 171–190 (1957).
4. E. T. Jaynes, *Probability Theory: The Logic of Science*, Cambridge Univ. Press, 2003.
5. P. C. Gregory, *Bayesian Logical Data Analysis for the Physical Sciences*, Cambridge Univ. Press, 2005.
6. D. G. Tarboton, R. L. Bras, and I. Rodriguez-Iturbe, *Water Resour. Res.* **25**, 2037–2051 (1989).
7. I. Rodriguez-Iturbe, E. J. Ijjasz-Vasquez, R. L. Bras, and D. G. Tarboton, *Water Resour. Res.* **28**, 1089–1093 (1992).
8. I. Rodriguez-Iturbe, and A. Rinaldo, *Fractal River Basins: Chance and Self-Organization*, Cambridge Univ. Press, 1997.
9. D. Veneziano, and J. D. Niemann, *Water Resour. Res.* **36**, 1923–1936 (2000).
10. C. D. Peters-Lidard, F. Pan, A. Y. Hsu, and P. E. O'Neill, *IEEE GRSS* **3**, 1297–1299 (2001).
11. B. B. Mandelbrot, *The Fractal Geometry of Nature*, W H Freeman and Co., 1983.
12. A. Arneodo, G. Grasseau, and M. Holschneider, *Phys. Rev. Lett.* **61**, 2281 (1988).
13. W. Newman, and D. Turcotte, *Geophys. J. Int.* **100**, 433–439 (1990).
14. M. Vergassola, and U. Frisch, *Physica D* **54**, 58–64 (1991).
15. U. Frisch, *Turbulence: The Legacy of A. N. Kolmogorov*, Cambridge Univ. Press, 1995.
16. A. Arneodo, *Wavelet Analysis of Fractals: from the Mathematical Concepts to Experimental Reality*, Oxford Univ. Press, 1996.
17. A. Turiel, and N. Parga, *Phys. Rev. Lett.* **85**, 3325 (2000).
18. B. Lashermes, E. Foufoula-Georgiou, and W. Dietrich, *Geophys. Res. Lett.* **34**, L23S04 (2007).
19. V. Nieves, and A. Turiel, *J. Marine Syst.* **77**, 482–494 (2009).
20. D. Schertzer, and S. Lovejoy, *Resolution Dependence and Multifractal in Remote Sensing and Geographical Information System*, McGill Univ. Press, 1996.
21. I. Daubechies, *Ten Lectures on Wavelets*, CBMS-NSF Series in App. Math., SIAM, 1992.
22. S. Mallat, *A Wavelet Tour of Signal Processing*, Academic, 1999.
23. C. E. Shannon, *Bell Syst. Tech. J.* **27**, 379–423 (1948).